The Two Size Alleles of Human Keratin 1 Are Due to a Deletion in the Glycine-Rich Carboxyl-Terminal V2 Subdomain

Bernhard P. Korge, John G. Compton, Peter M. Steinert, and Dietmar Mischke

The Skin Biology Branch (BPK, JGC, PMS), National Institute of Arthritis and Musculoskeletal and Skin Diseases, National Institutes of Health, Bethesda, Maryland, U.S.A., and Institute for Experimental Oncology and Transplantation Medicine (DM), University Medical Center Rudolf Virchow, The Free University of Berlin, Berlin, F.R.G.

Two size variants of the type II human keratin 1 protein chain, termed 1a and 1b, have been described previously. Using amplification of genomic DNA by the polymerase chain reaction and sequence analysis we show here that the difference between these two alleles is due to a deletion of 21 bp in sequences encoding the V2 subdomain. This deletion corresponds to an entire glycine loop of seven amino acids. Pedigree analysis showed that the alleles are inherited as normal Mendelian traits. No additional alleles were detected in a survey of 88 alleles from 44 unrelated individuals, and the allelic frequency of 1a and 1b was 0.61 and 0.39. To determine the molecular basis of inherited dermatoses it is preferable to perform genetic linkage studies utilizing candidate genes directly as polymorphic markers. The PCR-based keratin 1 alleles characterized here, together with previously described PCR-based size variants in the keratin 10 gene, provide useful markers for the keratin clusters on chromosome 12 and 17, respectively. *J Invest Dermatol 99:697–702, 1992*

The large family of keratins is composed of more than 30 protein chains that assemble into keratin intermediate filaments (KIF), which form a stable cytoskeleton in most epithelial cells. Keratins are subdivided into acidic type I and neutral-basic type II chains and are differentially expressed along the different pathways of epithelial cell development and differentiation. In most epithelial cells a characteristic pair of type I and type II keratins are co-expressed [1–5]. For example, basal epidermal cells express keratins 5 (type II) and 14 (type I), whereas cells committed to terminal differentiation express the keratins 1 (K1) (type II) and 10 (K10) (type I). Like all intermediate filament protein chains, keratins possess a central α-helical rod domain of conserved secondary structure. The rod domain is generally flanked by end domains that are arranged in characteristic patterns that permit their classification into subdomains. These include subdomains with a high degree of homology to each other within type I and type II chains (H subdomains), subdomains that consist of quasi peptide repeats (V subdomains), and subdomains that are highly charged (E subdomains) [3,6]. In the case of K1, the rod domain is flanked on the amino-terminal side by the E1, V1, and H1 subdomains, and on the carboxyl-terminal side by the H2, V2, and E2 subdomains (Fig 1). In contrast to most other keratins, the V1 and V2 subdomains of the K1 and K10 proteins are extremely glycine rich [7–10].

Recently, point mutations have been discovered in keratin genes expressed in the epidermis, which likely cause the autosomal dominant blistering skin diseases epidermolysis bullosa simplex (EBS) and epidermolytic hyperkeratosis (EHK). In the case of EBS, different point mutations were encountered in either the keratin 14 [11,12] or keratin 5 [13] chain that reside in the beginning, middle, or end of the rod domain, respectively. In EHK, ultrastructural analysis suggested that defective keratins might be responsible for the disease ([14–18] and references therein). Indeed, a leucine → proline substitution in the conserved H1-subdomain of the K1 chain has been identified as the likely molecular basis for the disease in one large Caucasian EHK family [19,20]. The locations of mutations identified in keratin genes of EBS and EHK patients have demonstrated that at least the beginning, a distinct region in the middle, and the end of the rod domain as well as the conserved H1-subdomain are apparently essential for the structural integrity of KIF [17,21]. In order to continue and extend genetic linkage studies to identify gene defects in other epidermal disorders that may involve keratins, it is necessary to find appropriate polymorphic markers close to or preferentially in the potential candidate genes for the disease.

Size polymorphism at the protein level has been described for the human keratins 1, 4, 5, and 10 [22–24]. For the human K10 chain an extensive size polymorphism was found in the carboxyl-terminal V2 subdomain [25]. The frequency of the numerous alleles makes K10 a useful marker for the type I keratin cluster on chromosome 17q, the only marker identified at this locus at this time. The K10 variations occur in the glycine-rich sequences that are arranged as quasi-repetitive peptides. We have proposed that these sequences form a new structural protein motif that we have termed the glycine loop [26]. This motif can be described by the form $x(y)_n$ where x is

Manuscript received August 21, 1992; accepted for publication September 16, 1992.

Reprint requests to: Dr. Steinert, Skin Biology Branch, National Institute of Arthritis and Musculoskeletal and Skin Diseases, NIH, Building 6, Room 427, Bethesda, MD 20892.

Abbreviations:

CEPH: Centre d'Etude du Polymorphisme Humain
EBS: epidermolysis bullosa simplex
EHK: epidermolytic hyperkeratosis
IF: intermediate filaments
KIF: keratin intermediate filaments
K1: keratin 1
K10: keratin 10
PCR: polymerase chain reaction
SDS-PAGE: sodium dodecylsulfate–polyacrylamide gel electrophoresis
VNTR: variable-number tandem repeats

0022-202X/92/$05.00 Copyright © 1992 by The Society for Investigative Dermatology, Inc.

usually an aromatic or rarely a long-chain aliphatic residue, γ is usually glycine but may also include polar residues, and n is highly variable with 1–35 residues in examples identified so far. The aromatic or long-chain aliphatic residues are likely to be stacked closely in a linear array and the interspersed glycine residues project away from this axis by forming a loop. A limited amount of sequence variation can apparently be tolerated. This is supported by the K10 data and by results found with human loricrin, the major cornified envelope protein of terminal differentiated epidermal cells, where multiple sequence variations in one of its glycine loops were detected [27].

We have identified two K1 alleles in the carboxyl-terminal glycine loops of K1, which we used in our genetic linkage study of EHK [19]. In this work we have extended the characterization of this K1 polymorphism. We intended to settle a controversy in the field as to whether the two described K1 protein alleles 1a and 1b represent two different alleles of K1 [23] or are due to post-translational modifications or protein degradation. Concurrently, we also wanted to know whether the 21-bp deletion identified in the glycine loops of the carboxyl-terminus of K1 [20] could account for the apparent molecular weight difference of 1 kDa detected for the two K1 protein alleles [23]. Furthermore, we wanted to see whether additional K1 alleles exist that have not been observed at the protein level [23]. The published gene sequences of the human K1 have shown some sequence variations but no net insertions or deletions [7,8], which therefore were unlikely to be responsible for the observed size variations of the K1 protein chain. In this article we have amplified genomic DNA by polymerase chain reaction (PCR) to resolve these questions.

MATERIALS AND METHODS

Keratin Isolation Keratin-enriched protein fractions were prepared from human interfollicular epidermis, solubilized in sample buffer [0.5 M Tris/HCl, pH 6.8, 2.1% sodium dodecylsulfate (SDS), 10% mercaptoethanol, 10% glycerol, and bromophenol-blue], and separated by one dimensional SDS-polyacrylamide gel electrophoresis (PAGE) according to Laemmli [28] as described [5,23]. For immunoblot analysis, proteins from unstained gels were transferred by semi-dry blotting onto nitrocellulose sheets [29] and incubated with the monoclonal antikeratin antibody K8.60 (Sigma, St. Louis) detecting K1 and K10 as described [5,23].

DNA Isolation High-molecular-weight genomic DNA was isolated according to standard procedures [30]. Briefly, tissue samples obtained from squamous cell carcinomas of the head and neck, normal interfollicular epidermis, and 12 whole foreskins from normal individuals were either ground to a fine powder under liquid nitrogen, or cut into small pieces, and digested with proteinase K (200 μg/ml) followed by phenol extraction and ethanol precipitation. Finally, the precipitated DNA was dissolved in water.

PCR Procedures Amplifications of genomic DNA were done using the following primers (+, sense strand; −, antisense strand): for the amino-terminal domain of the K1 gene, the primer BK22(+), starting one nucleotide before the start-codon, 5′-CATGAGTCGACAGTTTAGTTCCGGGTCTGGGTACCG-AAGT-3′, and primer 4b(−), located at the boundary of exon 1 and intron 1, 5′-GAGAGAGAAACTCACCTTTGTCAATG-3′; and for the carboxyl-terminal domain of the K1 gene, primer BK14S(+), located at the boundary of intron 8 and exon 9, 5′-CTTCCTCTTGCAGCTGTGAGCACAA-3′ and BK27(−), located at the end of exon 9, 5′-TGGTTACTCCGGAA-TAAGTGGTAGAAAC-3′. The PCR reactions were performed with about 400 ng genomic DNA per 100 μl reaction with 1.5 mM $MgCl_2$, 100 μM dNTPs, and 0.15 μM of each of the primers. Amplification of the amino-terminal domain of the K1 gene was for 33 cycles (95°C, 1 min; 62°C, 1 min; 72°C, 1.5 min) using a DNA Thermal Cycler (Perkin Elmer Cetus). The carboxyl-terminal domain of the K1 gene was amplified for four cycles (95°C, 1 min; 66°C, 1 min; 72°C, 1.5 min) followed by 30 cycles (95°C, 1 min; 64°C, 1 min; 72°C, 1.5 min). Amplifications were preceded by a "hot start": 5-min incubation of the DNA and primer mix at 96°C, equilibration at 80°C, followed by addition of the premixed dNTPs, PCR-buffer, $MgCl_2$, and Taq polymerase (AmpliTaq, Cetus) and then again incubated for 2 min at 95°C before the above-described cycling began. In the experiment analyzing the inheritance of the two alleles of the K1 gene in all members of one large Centre d'Etude du Polymorphisme Humain family (CEPH-family) [31] we added to the PCR reactions 1.5 pmol of 5′-end labeled primer BK27(+) (about 10000 cpm per tube). PCR products were separated on a 2–3% Nusieve 3:1 agarose gel (FMC, Rockland, ME) as described [30]. Agarose gels used for autoradiography were first dried and then exposed to x-ray film. Some of the PCR products were subcloned into the pGEM-3Z vector (Promega, Madison, WI) for DNA sequencing.

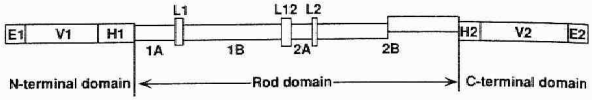

Figure 1. Domain organization of the human K1 protein chain. The α-helical rod domain with its subdomains 1A, 1B, 2A, and 2B and the linker regions L1, L12, and L2 is flanked on both sides by end domains. The amino-terminal domain consists of the E1, V1, and H1 subdomains; the carboxyl-terminal domain includes the H2, V2, and E2 subdomains. The two allelic variants of K1 were found in the V2 subdomain *(stippled box)*.

Statistical Tests χ^2-tests with Yates corrections were used to assess differences in allele frequencies between groups.

RESULTS

Size Variation in the Carboxyl-Terminal Domain of Two Human Keratin K1 Gene Clones The K1 chain possesses glycine loop motifs in the V1 and V2 subdomains (Fig 1). Earlier work on keratin variants in human epidermis identified two size alleles for the human K1 chain, termed 1a and 1b, differing in apparent molecular weight of 1 kDa [23]. More recently, we found by PCR a size variation in exon 9 of the human K1 gene in a single family [19]. We set out to establish whether the described alleles 1a and 1b can be accounted for by a PCR-detectable deletion in the V2 subdomain of K1 or may be due to deletions in the also glycine-rich V1 subdomain. Accordingly, PCR primers were designed to amplify both the V1 and V2 subdomains of K1. To define appropriate PCR conditions we used two previously isolated human genomic K1 clones, H67 [8] and 4D1a [Bowden et al (abstr). J Invest Dermatol 88:478A, 1987]. We found that although the PCR products encoding for the V1 subdomain revealed no size variations for these two genomic clones (Fig 2, *lanes 1* and *2*), there was a considerable size difference in the PCR products encoding for the V2 subdomain of the K1 gene (Fig 3, *lanes 1* and *2*). Thus, these two human genomic K1 clones represent two different alleles.

Direct Correlation of K1 Alleles on the Protein and Gene Level We analyzed five unrelated individuals for the presence of the 1a and 1b size alleles in their expressed K1 protein chains and in their genomic DNA. SDS-polyacrylamide gels separating keratin-enriched fractions were either stained with Coomassie-Blue or blotted on a nitrocellulose filter for immunodetection. In both analyses two bands of different size representing the two alleles, 1a and 1b, were identified (Fig 4*a*,*b*). Then, using genomic DNA from these same individuals, the V1 and V2 subdomains were amplified by the same PCR technique as described above. The PCR products for the V1 subdomain again did not show any size variations (data not shown). However, for the V2 subdomain the PCR products displayed either a single band or two bands of equal intensity (Fig 4*c*). For example, individual 2 showed on the protein gel and immunoblot (Fig 4*a*,*b*) as well as on the DNA gel (Fig 4*c*) a single band corresponding to the previously described lower molecular weight

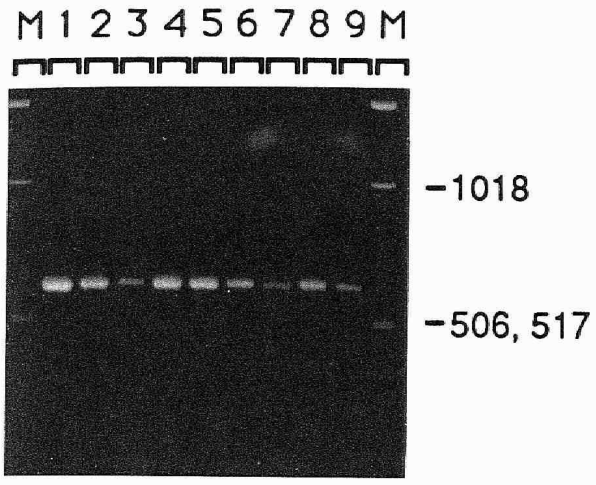

Figure 2. PCR amplification of the gene sequence encoding the amino-terminal domain of human K1. Two genomic human K1 clones and 7 genomic DNA samples from unrelated individuals were amplified by PCR and electrophoretically separated on a 2% agarose gel. *Lanes 1* and *2* show the amplified PCR products of the two genomic human K1 clones H67 and 4D1a, respectively. *Lanes 3* through *9* show PCR products for the seven genomic DNA samples. The DNA size marker was the 1-kilobase DNA ladder (GIBCO/BRL, Gaithersburg).

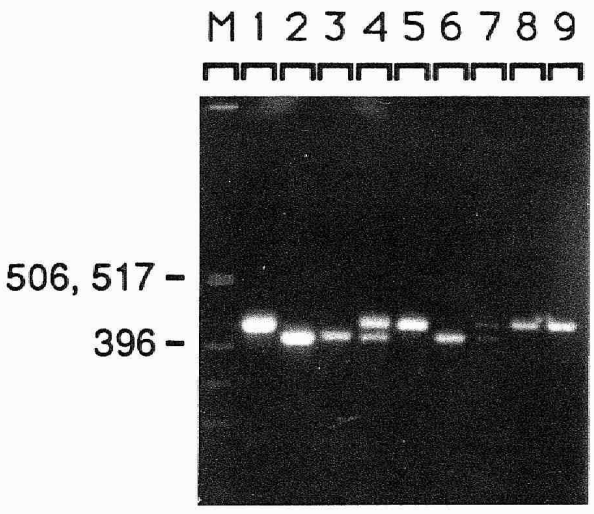

Figure 3. PCR amplification of the gene sequence encoding the carboxyl-terminal domain of human K1. DNA of the same two genomic K1 clones and same seven genomic DNA samples as in Fig 2 were used for amplification of the carboxyl-terminus of K1. PCR products were electrophoretically separated on a 3% agarose gel. *Lanes 1* and *2* show the PCR products of the two genomic human K1 clones H67 and 4D1a, respectively, and *lanes 3* through *9* show PCR products for the seven genomic DNA samples. The DNA size marker was the 1-kilobase DNA ladder (GIBCO/BRL, Gaithersburg).

1b allele of the K1 protein (Mischke and Wild, 1987). Thus, individual 2 is homozygous for the 1b allele. Individuals 4 and 5 also showed single, but larger, bands on the protein (Fig 4*a*,*b*) and DNA (Fig 4*c*) gels, consistent with the higher-molecular-weight 1a allele of the K1 protein. Thus individuals 4 and 5 are homozygous for the larger 1a allele. Analyses of individuals 1 and 3 displayed two bands on the protein (Fig 4*a*,*b*) and DNA (Fig 4*c*) gels corresponding in size to the 1a and 1b alleles and are therefore heterozygous for both the 1a and 1b alleles. Hence, there is an exact correlation between the size alleles identified for the K1 protein and the PCR products obtained from the gene sequences encoding for the V2 subdomain of the same individuals.

Sequence Analysis of the Two K1 Alleles To elucidate the nature of the variations in the carboxyl-terminal domain, we sequenced our two genomic clones as well as one homozygous and one heterozygous individual (a total of six different alleles). In each case the 1b allele had a 21-bp deletion in one location of the V2 subdomain with respect to the 1a allele. This 21-bp encoded the seven amino acids Gly-Ser-Gly-Gly-Ser-Ser-Tyr resulting in the deletion of an entire glycine loop (Fig 5). We have previously encountered, because of heteroduplex formation, two different alleles of the same size resulting from the same net deletion on two different locations in the loricrin gene [27]. Therefore we analyzed the apparently same-sized PCR products of the V2 subdomain of K1 on non-denaturing polyacrylamide gels to detect heteroduplex formation in an attempt to identify additional K1 alleles. Heteroduplex formation occurs during PCR amplification if two different single-stranded alleles hybridize to each other. These heteroduplexes differ in their electrophoretic mobility from homoduplexes in a manner analogous to single-strand conformation analysis [32]. In 16 unrelated individuals analyzed, heteroduplex formation was only observed in the four heterozygous individuals but in none of the 12 homozygous individuals (data not shown). Because alleles that are equal in size but slightly different in sequence would still form detectable heteroduplexes, their absence in the analyzed individuals established the presence of only a single deletion site.

Survey for Other-Size Alleles of the K1 Gene Within the Human Population Using the same specific primers and PCR conditions as defined above, a larger number of unrelated individuals were screened for the presence of size variations in the V1 and V2 subdomains. For the V1 subdomain a total of 38 alleles were analyzed. No size variants could be detected in any of them (a selection of seven individuals is shown in Fig 2, *lanes 3–9*). Analyses of the V2 subdomain of 88 allels from 44 unrelated individuals rendered only the same two alleles 1a and 1b identified above (the PCR products of the same individuals as shown in Fig 2 are shown for the V2 subdomain in Fig 3, *lanes 3–9*). In all cases only one band or two bands of equal intensity were found, suggesting that the different-sized PCR products represent two alleles of a single-copy gene per haploid genome.

Inheritance of the K1 Sequence Polymorphism In order to demonstrate the inheritance pathway of the K1 polymorphism, we amplified the V2 subdomain of one kindred CEPH family [31]. The 12 children were either homozygous for the 1a allele or heterozygous for the 1a and 1b alleles present in the parents (Fig 6). Together these data clearly demonstrate that the size polymorphism of the K1 gene segregates as a normal Mendelian trait.

Allele Frequency of the Two K1 Alleles We have surveyed a total of 44 unrelated individuals of different ethnic origin. We observed 16 individuals homozygous for the 1a allele, six individuals homozygous for the 1b allele, and 22 heterozygous individuals. The calculated allele frequency was 0.61 (95% confidence interval is 0.51–0.71) for the 1a allele and 0.39 (95% confidence interval is 0.29–0.49) for the 1b allele. The expected distribution of the two alleles in the population according to the Hardy-Weinberg law was consistent with the observed frequency. The allele frequencies in our population are statistically different ($p = 0.02$) from the published allele frequencies of 0.47 for the 1a allele and 0.53 for the 1b allele obtained for 148 individuals at the protein level [23]. This

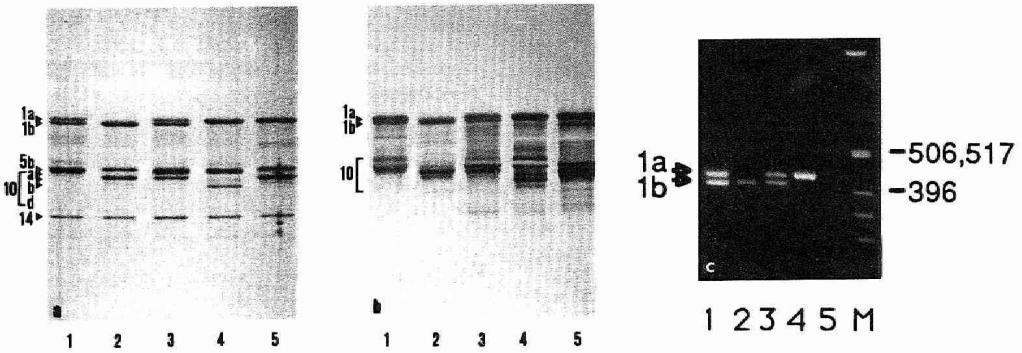

Figure 4. Direct correlation of K1 alleles on the protein and gene level of K1. *(A)* Coomassie-Blue stained gel (SDS/PAGE) of epidermal keratins from five unrelated individuals. *(B)* Corresponding immunorecognition pattern using the monoclonal antikeratin antibody K8.60 (Sigma, St. Louis) detecting K1 and K10. *(C)* 3% agarose gel showing PCR products from the amplified gene sequence encoding the carboxyl-terminal domain of K1 using genomic DNA from the same individuals. The patterns shown in *A* and *B* display the kertains [1a + 1b] + 5b + 10a + 14 *(lane 1)*; 1b + 5b + 10b + 14 *(lane 2)*; [1a + 1b] + 5b + [10a + 10b] + 14 *(lane 3)*; 1a + 5b + [10a + 10d] + 14 *(lane 4)*; and 1a + 5b + 10b + 14 *(lane 5)*. Note that keratins 5b and 10a co-migrate under these conditions, but may be distinguished by immunoblotting with appropriated antibodies. The PCR products shown in *C* reveal that the number of bands and the size variations found correspond with changes found for the K1 proteins in *A* and *B*. The molecular weight maker (M) used in *C* is the 1-kilobase DNA ladder (GIBCO/BRL, Bethesda).

difference is most likely due to the analysis of two quite different ethnic populations.

DISCUSSION

Previously we have found two allelic variants for the K1 protein, 1a and 1b, differing in apparent molecular weight [23]. Here we characterize two allelic variants in the K1 gene in which the size of the V2 subdomain differs. Furthermore, we demonstrate that the appearance of alleles 1a and 1b of K1 protein chain corresponds to the presence of these two K1 gene variants, and that the lower-molecular-weight allele is due to the deletion of an entire glycine loop of seven amino acids in the V2 subdomain. These findings unambiguously establish that the described polymorphism of the K1 protein chain does indeed represent two different alleles and is not due to post-translational modifications or protein processing in the course of sample preparation. Similar variations have been demonstrated for the V2 subdomain of K10 and in one of the glycine loops of the human loricrin.

We have shown that the glycine loop motif is highly flexible [33] and probably responsible, at least in part, for the maintenance of stratum corneum flexibility and elasticity, due to interactions of the glycine loops on KIF and loricrin of the cell envelope (the "Velcro hypothesis") [26]. Our data on the sequence variations of K10 and loricrin, and now of K1, reveal that a limited degree of sequence divergence does not alter the structural organization or function of these motifs.

Why do the size variations in the K1 and K10 chains occur preferentially in the carboxyl-terminal V2 subdomain rather than the amino-terminal V1 subdomain? Several types of studies have shown that the structural integrity of the amino-terminus is far more important for intermediate filament (IF) assembly and function than the carboxyl-terminus. For example, previous studies have shown that *in vitro* assembly of IF is not impaired using carboxyl-terminal truncated forms of the desmin [34], vimentin [35], keratin 8, or keratin 18 [36] chains. Transfection experiments with amino- and carboxyl-terminal deletion constructs of K14 into simple epithelial cells and squamous cell carcinomas [37–39] demonstrated that IF formation is more disturbed by deletion of the non-helical amino-terminus than of the carboxyl-terminus. Furthermore, assembly studies with vimentin demonstrated the importance of the amino-terminus in stabilizing IFs [40,41] and revealed that the spacing of a specific nonapeptide at the amino-terminus with regard to the rod domain is crucial for IF stability [41]. Our present data are consistent with the possibility that the integrity of the glycine loops in the amino-terminus of the K1 and K10 chains also may be important for normal KIF assembly, structure, and function.

Why is the V2 subdomain of K10 far more polymorphic than that of K1? Although we have no direct data, a possible explanation may exist in the observation that the nucleic acid sequences encoding the quasi-peptide repeats of K10 are more precisely conserved than those for K1. Studies with other systems have shown that exact tandem repeats favor the occurrence of unequal crossing-over events during meiosis as described for rRNA genes [42,43] and many multigene families (for review see [44]). The mechanism of unequal crossing-over was earlier proposed for the generation and expansion of glycine-rich sequences in keratins [45,46]. Alternatively, Wolff et al [47] suggest replication slippage as a likely mechanism for the generation of new-length alleles in VNTRs (variable-number tandem repeat) and reject the unequal crossing-over model as the major mechanism. Nevertheless, because there is greater sequence identity in the tandem quasi-repeats of the V2 subdomain of K10 compared to K1, either mechanism could account for the more extensive occurrence of size variants in K10.

The polymorphism of the K1 gene is extremely useful for the continuation and extension of genetic linkage work with inherited skin diseases in which genes of the type II keratin cluster on chromosome 12q may or may not be involved. A general genome search in order to link an inherited disorder to a distinct chromosomal locus involves an enormous amount of work. Therefore, as a preferred strategy, attempts are made to start out with potential candidate genes that could be responsible for a given disease phenotype. For this, polymorphic markers very near or even in the candidate gene are significantly more informative, especially for exclusion, than linked markers at unknown physical distance(s). The absence of recombination leads to a high lod score improving the likelihood of linkage. However, if a marker is genetically even a few centi-Morgans (1 centiMorgan is a statistical measurement of genetic distance, which is about 1 Mb in the human genome) from the disease locus apart, then a recombination event may not necessarily exclude linkage. In two recent studies of EBS and EHK [11,19], genetic linkage analysis was used to establish linkage to the type I or type II keratin cluster. Polymorphic markers near and in the type I keratin cluster on chromosome 17 were utilized. However, similar polymorphic markers near or within the type II keratin cluster on

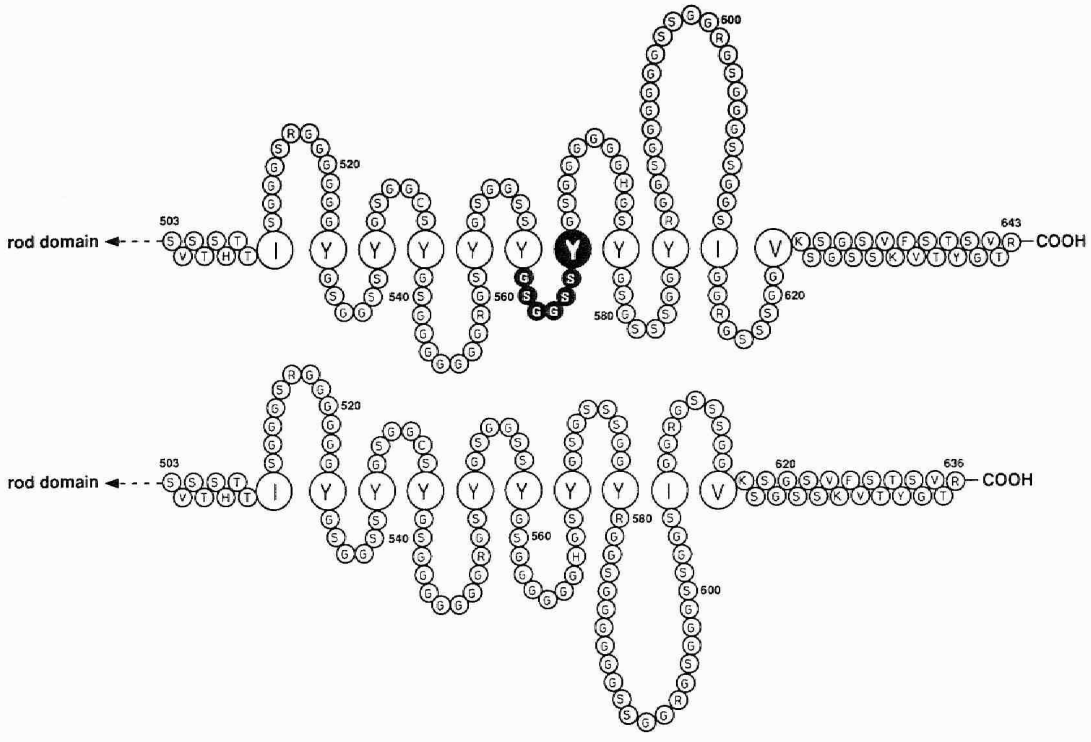

Figure 5. Two-dimensional model for the two K1 alleles in the V2 subdomain. This drawing shows a two-dimensional representation of the glycine loop motif present in the V2 subdomain of the 1a and 1b allele of human K1 based on our recent model for such sequences [26]. The region deleted in the smaller 1b allele with respect to the larger 1a allele is highlighted in reverse contrast.

chromosome 12 had not been described, until the presently analyzed polymorphism in the glycine loop sequences of K1 was identified. Thus the availability of this well-characterized polymorphic marker in the K1 gene within the type II cluster is of enormous benefit for future linkage work involving keratins as potential candidate genes.

We wish to express our thanks to Drs. Constantin Chipev, Chris Amos, Sherri Bale, and Wesley McBride for their helpful suggestions, contributions, and comments during the course of this work. We thank Gabriele Wille for her expert technical assistance with the protein gels and immunoblots. Will Idler isolated both genomic K1 clones. George Poy kindly synthesized all of the oligonucleotides used in this work. Bernhard P. Korge and Dietmar Mischke both acknowledge support from the Deutsche Forschungsgemeinschaft.

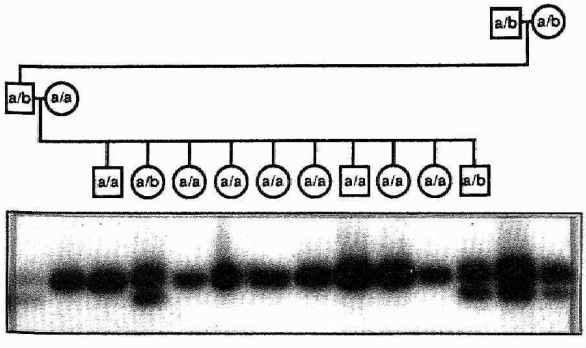

Figure 6. Mendelian segregation of the two human K1 size alleles. DNA from transformed lymphocytes of a three-generation kindred family (CEPH cell lines) [31] was used for PCR amplification of the carboxyl-terminal end domain of K1 as before. Both the autoradiogram (3% agarose gel) and pedigree distribution are shown. The pedigree is aligned to the gel lanes. The two different alleles are denoted as 1a and 1b according to the earlier suggested nomenclature [23].

REFERENCES

1. Moll R, Franke WW, Schiller DI, Geiger B, Krepler R: The catalog of human cytokeratins: patterns of expression in normal epithelia, tumors and cultured cells. Cell 31:1–24, 1982
2. Dale BA, Holbrook KA, Kimball JR, Hoff M, Sun T-T: Expression of epidermal keratins and filaggrin during fetal skin development. J Cell Biol 101:1257–1269, 1985
3. Steinert PM, Roop DR: The molecular and cellular biology of intermediate filaments. Annu Rev Biochem 57:593–625, 1988
4. O'Guin WM, Schermer A, Lynch M, Sun T-T: Differentiation-specific expression of keratin pairs. In: Goldman RD, Steinert PM (eds.). Cellular and Molecular Biology of Intermediate Filaments. Plenum Publishing Corporation, New York, 1990, pp 301–334
5. Korge B, Stadler R, Mischke D: Effect of retinoids on hyperproliferation-associated keratins K6 and K16 in cultured human keratinocytes: a quantitative analysis. J Invest Dermatol 95:450–455, 1990
6. Conway JF, Parry DAD: Intermediate filament structure: 3. Analysis of sequence homologies. Int J Biol Macromol 10:79–98, 1988
7. Steinert PM, Parry DAD, Idler WW, Johnson LD, Steven AC, Roop DR: Amino acid sequences of mouse and human epidermal type II

keratins of 67,000 molecular weight provide a systematic basis for the structure and functional diversity of the end domains of keratin intermediate filament subunits. J Biol Chem 260:7142–7149, 1985

8. Johnson LD, Idler WW, Zhou X-M, Roop DR, Steinert PM: Structure of a gene for the human epidermal keratin of 67,000 Da. Proc Natl Acad Sci USA 82:1896–1900, 1985
9. Zhou X-M, Idler WW, Steven AC, Roop DR, Steinert PM: The sequence and structure of human keratin 10: organization and possible structures of end domain sequences. J Biol Chem 263:15584–15589, 1988
10. Rieger M, Franke WW: Identification of an orthologous mammalian cytokeratin gene. High degree of intron conservation during evolution of human cytokeratin 10. J Mol Biol 204:841–856, 1988
11. Bonifas JM, Rothman AL, Epstein EH: Epidermolysis bullosa simplex: evidence in two families for keratin gene abnormalities. Science 254:1202–1205, 1991
12. Coulombe PA, Hutton ME, Letai A, Hebert A, Paller AS, Fuchs E: Point mutations in human keratin 14 genes of epidermolysis bullosa simplex patients: genetic and functional analyses. Cell 66:1301–1311, 1991
13. Lane EB, Rugg EL, Navsaria H, Leigh IM, Heagerty AHM, Ishida-Yamamoto A, Eady RAJ: A mutation in the conserved helix termination peptide of keratin 5 in hereditary skin blistering. Nature 356:244–246, 1992
14. Ogawa H, Hattori M, Ishibashi Y: Abnormal fibrous protein isolated from the stratum corneum of a patient with bullous congenital ichthyosiform erythroderma (BCIE). Arch Dermatol Res 266:109–116, 1979
15. Anton-Lamprecht I: Genetically induced abnormalities of epidermal differentiation and ultrastructure in ichthyoses and epidermolyses; pathogenesis, heterogeneity, fetal manifestation, and prenatal diagnosis. J Invest Dermatol 81:149S–156S, 1983
16. Resing KA, Dale BA: Proteins of keratohyalin. In: Goldsmith LA (ed.). Physiology, Biochemistry and Molecular Biology of the Skin. Oxford University Press, New York, 1991, pp 148–167
17. Fuchs E, Coulombe PA: Of mice and men: genetic skin diseases of keratin. Cell 69:899–902, 1992
18. Ishida-Yamamoto A, McGrath JA, Judge MR, Leigh IM, Lane EB, Eady RAJ: Selective involvement of keratins K1 and K10 in the cytoskeletal abnormality of epidermolytic hyperkeratosis (bullous congenital ichthyosiform erythroderma). J Invest Dermatol 99:19–26, 1992
19. Compton JG, DiGiovanna JJ, Santucci SK, Kerans KS, Amos CL, Abangan DL, Korge BP, McBride OW, Steinert PM, Bale SJ: Linkage of epidermolytic hyperkeratosis to the type II keratin gene cluster on chromosome 12q. Nature genetics 1:301–305, 1992
20. Chipev CC, Korge BP, Markova N, Bale SJ, DiGiovanna JJ, Compton JG, Steinert PM: A leucine→proline mutation in the H1 subdomain of keratin 1 causes epidermolytic hyperkeratosis. Cell 70:821–828, 1992
21. Parry DAD, Steinert PM: Intermediate filament structure. Curr Opin Cell Biol 4:94–98, 1992
22. Wild G-A, Mischke D: Variation and frequency of cytokeratin polypeptide patterns in human squamous non-keratinizing epithelium. Exp Cell Res 162:114–126, 1986
23. Mischke D, Wild G: Polymorphic keratins in human epidermis. J Invest Dermatol 88:191–197, 1987
24. Mischke D, Wille G, Wild GA: Allele frequencies and segregation of human polymorphic keratins K4 and K5. Am J Hum Genet 46:548–552, 1990
25. Korge BP, Gan SQ, McBride OW, Mischke D, Steinert PM: Extensive size polymorphism of the human keratin 10 chain resides in the carboxyl-terminal V2-subdomain due to variable numbers and sizes of glycine loops. Proc Natl Acad Sci USA 89:910–914, 1992
26. Steinert PM, Mack JW, Korge B, Gan S-Q, Haynes SR, Steven AC: Glycine loops in proteins. Their occurrence in certain intermediate filament chains, loricrins and single-stranded RNA binding proteins. Int J Biol Macromol 13:130–139, 1991
27. Yoneda K, Hohl D, McBride OW, Wang M, Cehrs KU, Idler WW, Steinert PM: The human loricrin gene. J Biol Chem 267:18060–18066, 1992
28. Laemmli UK: Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 277:680–685, 1970
29. Towbin H, Staehelin T, Gordon J: Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedures and some applications. Proc Natl Acad Sci USA 76:4350–4354, 1979
30. Sambrook J, Fritsch EF, Maniatis T: In: Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1989
31. White RL, Lalouel J-M, Nakamura Y, Daus-Keller H, Green P, Bowden DW, Mathew CGP, Easton DF, Robson EB, Morton NE, Gusella JF, Haines JL, Retief AE, Kidd KK, Murray JC, Lathrop GM, Cann HM: The CEPH consortium primary linkage map of human chromosome 10. Genomics 6:393–412, 1990
32. Orita M, Suzuki Y, Sekiya T, Hayashi K: Rapid and sensitive detection of point mutations and DNA polymorphisms using polymerase chain reaction. Genomics 5:874–879, 1989
33. Mack JW, Torchia DA, Steinert PM: Solid state NMR studies on the dynamic structure of mouse epidermal keratin filaments. Biochemistry 27:5418–5426, 1988
34. Kaufmann E, Weber K, Geisler N: Intermediate filament forming ability of desmin derivatives lacking either the amino-terminal 67 or the carboxyl-terminal 27 residues. J Mol Biol 185:733–742, 1985
35. Shoeman Rl, Mothes E, Kesselmeier C, Traub P: Intermediate filament assembly and stability in vitro: effect and implications of the removal of head and tail domains of vimentin by the human immunodeficiency virus type I protease. Cell Biol Int Rep 14:583–594, 1990
36. Hatzfeld M, Weber K: Tailless keratins assemble into regular intermediate filaments in vitro. J Cell Sci 97:317–324, 1990
37. Albers K, Fuchs E: The expression of mutant epidermal keratin cDNAs transfected in simple epithelial and squamous cell carcinoma lines. J Cell Biol 105:791–806, 1987
38. Albers K, Fuchs E: Expression of mutant keratin cDNAs in epithelial cells reveals possible mechanisms for initiation and assembly of intermediate filaments. J Cel Biol 108:1477–1493, 1989
39. Coulombe PA, Chan Y-M, Albers K, Fuchs E: Deletions in epidermal keratins leading to alterations in filament organization in vivo and in intermediate filament assembly in vitro. J Cell Biol 111:3049–3064, 1990
40. Traub P, Scherbarth A, Wiegers W, Shoeman RL: Salt-stable interaction of the amino-terminal head region of vimentin with the a-helical rod domain of cytoplasmic intermediate filament proteins and its relevance to protofilament structure and filament formation and stability. J Cell Sci 101:368–381, 1992
41. Hermann H, Hoffmann I, Franke WW: Identification of a nonapeptide motif in the vimentin head domain involved in intermediate filament assembly. J Mol Biol 223:637–650, 1992
42. Petes TD: Unequal meiotic recombination within tandem arrays of yeast ribosomal DNA genes. Cell 19:765–774, 1980
43. Petes TD, Hill CW: Recombination between repeated genes in microorganisms. Annu Rev Genet 22:168–174, 1988
44. Ohta T: On the evolution of multigene families. Theor Popul Biol 23:216–240, 1983
45. Klinge EM, Sylvestre YR, Freedberg IM, Blumenberg M: Evolution of keratin genes: different protein domains evolve by different pathways. J Mol Evol 24:319–329, 1987
46. Steinert PM, Roop DR: The structure, complexity and evolution of intermediate filament genes. In: Goldman RD, Steinert PM (eds.). Cellular and Molecular Biology of Intermediate Filaments. Plenum Publishing Corporation, New York, 1990, pp 353–367
47. Wolf RK, Plaetke R, Jeffreys AJ, White R: Unequal crossingover between homologous chromosomes is not the major mechanism involved in the generation of new alleles at VNTR loci. Genomics 5:382–384, 1989